Volume 11, Special Issue I, May 2020
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
CONFERENCE PAPER
Available Online at www.ijarcs.info

MACHINE LEARNING-BASED WEATHERFORECASTING ON FREELY AVAILABLE WEATHER DATA

Karthik Sudershan
School of Computer Science and Engineering
REVA University, Bengaluru, India
Email-karthiksudarshan7117@gmail.com

Vishnu Soman
School of Computer Science and Engineering
REVA University, Bengaluru, India
Email-s.vishnusoman33@gmail.com

Kiran K
School of Computer Science and Engineering
REVA University, Bengaluru, India
Email-kiran1994123@gmail.com

Shreenidhi S Deshpande
School of Computer Science and Engineering
REVA University, Bengaluru, India
Email- shreenidhideshpande4u@gmail.com

Prof. Surendra Babu KN
REVA University Bangaluru, India
Email-surendrababukn@reva.edu.in

Abstract- **Weather forecasting is mostly a numerical measurerather than a binary conclusion. Precise weather forecasting is one of today's highlyinteresting tasks which deals with anenormousamount of observations and features. We aimto develop an intellectual weather forecasting module. Asmartforecast based on the freely available data is achievedwith machine learning techniques. This module reflects trails such as maximum temperature, minimum temperature, and rainfall for an experimented period of years, and they are evaluated. The dataset is repossessed from a website named DarkSky.To abridge the recovery of data,a Python API is used to studymeteorological data has been established. The study is grounded on Linear Regression and anArtificial Neural Network model thatforecast next day weather with goodprecision. Anaccurateness of more than 94% is gained based on the dataset. Newexplorations haveshown that ML methods accomplishedbetter presentation than old-style statistical approaches. Machine learning, a subdivision of artificial intelligence has establishedto be a robust method offoreseeing and examining a given data set. This unit plays asignificantpart in the field of agriculture where weather forecasting is a vital aspect.**

Keywords- Machine learning; Weather Predicting; Linear Regression; Artificial Neural Network.

I. INTRODUCTION

Weather predictionis a process that is a challenging and very complex task to do so. This is because it needs various records and data of our Environment to forecast the weather. Weather prediction is used for Agriculture, disaster management, Travelling,etc. which plays a major role for them. Therefore, the forecasting of the weather plays a very important and major role in our day to day lives. Businesses across the globe are dependent on weather forecasting to sustain their trades.

This is one of the major reasons why weather forecasting is not a simple process.In current times there are satellites that capture high-resolution images to accurately find the forecast ofweather for the upcoming days, but this process is not simple and also not cheap.

There has been a lot of progress in the science of meteorology to date. Now, meteorologists use methods that were proposed by Lewis Fry Richardson in the year 1922. But weather prediction began much later, in the year 1955, which was made possible by the development of programmable electronic computers.This electronic computer always kept note metrics as data such as wind speed, temperature, wind direction, precipitation atmospheric pressure, and humidity.

Traditional methods include physical simulations in which atmosphere are modeled as fluid. But the original equations that are used in this physical model is unstable, and the initial metrics are not certain to us which are obtained by or from the atmosphere. This in turn is something that is not fully completed to understand the complexity of the atmosphere and it also creates a restraint to predict the weather accurately.

II. LITERATURE SURVEY

For over well over a decade, many scholars have been exploring ways to improve the current methods of the weather forecast. When data is abundant, climate scientists build many data-driven models. These models are generally constructed by explaining regression and classification problems, and new ML techniques can solve numerous complications that were previously stimulating [1]. Python APIs have been advanced that abridge the recovery and administration of meteorological data confined in large data sets online. Apart from modest ML models that employ regression, even more progressive constructions have been planned for solving the problem with the autoregressive structure with input (ARX).

In such a model, the present (local) set of properties is interrelated with both previous values of the similar (local) set, but also with other standards from the same position or standards of the same properties from other locations at anexisting time [2]. ACM models deliver an account of numerous climatic variables (temperature, humidity, geopotential, wind components, etc.) that define the foreseen atmospheric design for a given forecast period [4].

This paper demonstrates how Bayesian Networks (BNs) offer a sound and applied method for this problem, permitting to inevitably construct submissive probabilistic models from data determining the current dependencies among the positions within the databases [4]. They use the MLP network model which foresees the temperature of a place a day ahead and this network has a good recital and sensible forecast accurateness attained for this model. It's predicting reliabilities were estimated by calculating the mean absolute error between the precise and foreseen values [5]. For forecasting monsoon rainfall in districts and subdivision areas for about a week or 10 days using dynamical atmospheric universal circulation models definite boundary conditions and changing initial conditions [6]. In this paper, they discover mechanically
generating site-specific prediction models for solar power generation from NationalWeather Service (NWS) weather predictions using a machine learning method [7].

They use Classification and Regression Tree algorithm, Naive Bayes approach, K means Neighbours and Neural Networking for predicting Rainfall and its exactness is verified on a test dataset [8]. In this paper, they propose to find the minimum and maximum temperature and rainfall calculation of the next day by Artificial Intelligence [9].

This paper uses linear regression and support vector machine methods of machine learning and then use augmented algorithm rule for foreseeing the weather for the next 5 days.

To get a more precise outcome, they compute it with arithmeticaldecision tree and conditions vide confusion matrix using Big data [10]. This paper uses the Soil Moisture and Ocean Salinity data to analyse the soil moisture for the Numerical Weather Forecast application. These two are advanced grounded on SMOS brightness temperature and SMOS neural network soil moisture data assimilation [11].In this paper, they use machine learning to foresee weather with the Dynamic Integrated estimating system which is the first automated weather forecasting engine [12].

This paper is a two-step rationalizing for the GCM forecast. In the first step, it uses cluster analysis for recognizing the circulation design. In the second step, a conditional stepwise screening regression analysis is achieved for each weather element and other areas [13]. This paper is based on a neural architecture that combines a self-organizing map (SOFM) and MLPs to understand a hybrid network called SOFM–MLP. Then it combines with the FSMLP to get an improved outcome with very limited inputs and gives virtuous forecast [14]. In this paper, a black-box modeling technique is proposed for temperature predicting. Due to the high dimensionality of data, feature selection is completed in binary steps with k-Nearest Neighbours and Elastic net. Then the forecasts of the model were equated with weather underground forecasts and it displays that they both are comparable [16].

In this paper, they use one-hour-ahead load predicting technique using the improvement of similar day data. Then the estimated load power is gained by adding a modification to the particular similar day data [17]. This paper has two key workings: weather-traffic index establishment and key factor analysis. Using this system, they showed a complete practical study in a city, and the weather-traffic indices removed have been confirmed to be amazingly consistent with real-world observations. Additional regional key factor examination produces stimulating outcomes [18].

III. METHODOLOGY

1. Data collection and processing - For simplification and ease of access, we will use APIs to navigate the large sets of meteorological data, based on which we will make predictions. Once collected, it will be processed and aggregated into a format that is suitable for data analysis, then cleaned. 2. Linear Regression Models. Performing analysis on the data set so as to choose beneficial features using stats and scikit-learning libraries in Python. 3. Neural Network Models- Designing an ANN model and inferring the relationship in results between the two models.

In Linear Regression, a set of initial assumptions wrt linear relationships and numerical methods to find an outcome, which is based multiple predictors.

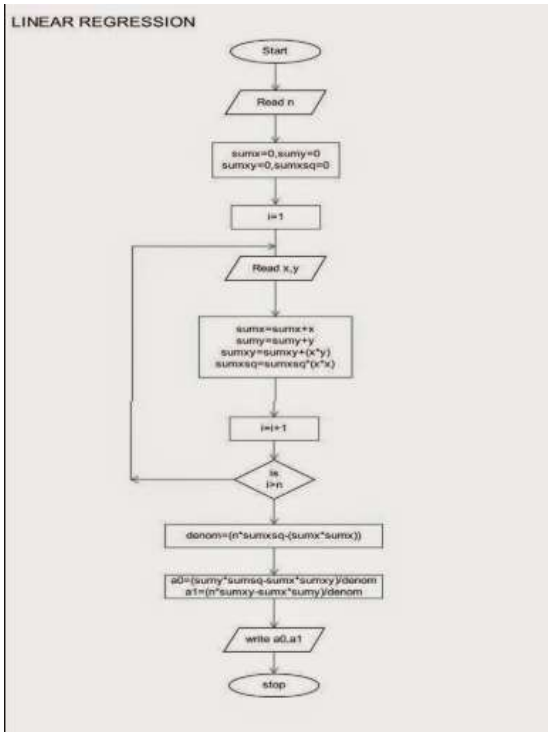
Linear Regression model is as follows:

$$\hat{y} = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \ldots + \beta_{(p-n)}\, x_{(p-n)} + E$$

where:

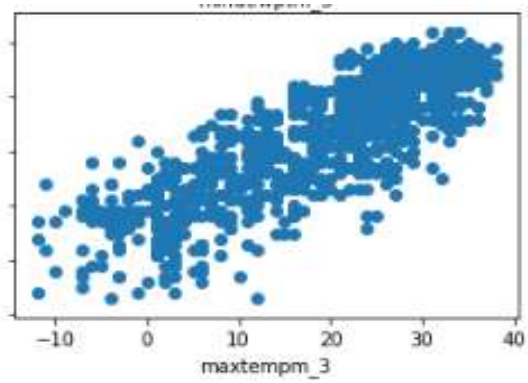
- $\hat{y}$ being the outcome
- x_j being the predictor variables
- β_0 being the intercept
- β_i being the change in $\hat{y}$
- E being a random error term

Linear regression needs a key assumption. One method of assessing the linearity between our independent variable, the mean temperature, and the other independent variables is to calculate the Pearson correlation coefficient.



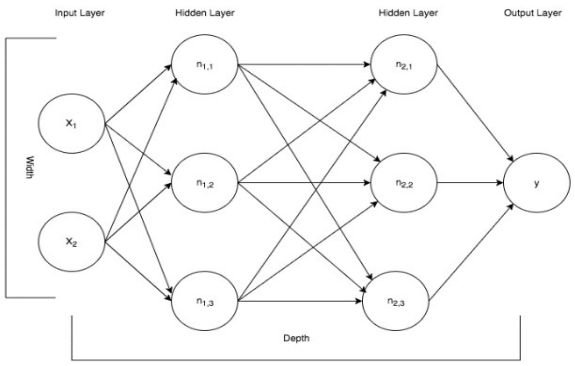
The Pearson correlation coefficient (r) is the measureof the amount of linear correlation between arrays of equal lengthis

outputted as a value ranging -1 to 1. These values between 0 and 1 represent a progressively positive correlation.



Neural networks areinspired by the human brain. They have nearly endless applications. An ANN is similar to the brain in a manner that they both have nodes and are a collection of neurons. The input is forwarded through each such agent and is an important feature of an ANN.

If represented graphically, a neural network similar to the one being described in this paper is shown in the imagebelow:



The ANN shown above has an input layer with two inputs. It has two hidden layers with 3 nodes each. Finally, an output layer that outputs one numerical value using a cost function. In this case the sum of squared errors.

Model optimization algorithms are crucial in designing robust neural networks. As examples are supplied through the network's architecturethen evaluated w.r.t the cost function, the weights are adjusted. When the optimizer function notices that a weight adjustment was made in such a way that does not improve the cost function, the model is said to be "learning", which is then saved with the optimizer so that it does not adjust the weights in that direction again

IV. CONCLUSION

This project has been to demonstrate the use of a DNNRegressor.To demonstrate this process of predicting the weather we have created an ANN model capable of predicting the next day's weather. This has largely been done on the features selected from freely available weather data. We have achieved goals as follows :

- Demonstrate the general process for undertaking ananalytics project.
- Show how to select meaningful features that do not violate key assumptions of the Linear Regression.
- Demonstrate how theuse of high-level TensorFlow API
- Talk about the problems relatedto overfitting a model.
- Explain the importance of experimenting with more than one model type to best solve a problem.

REFERENCES

[1] David Rolnick,Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont9, Natasha Jaques, Anna Waldman-Brown, Alexandra Luccioni6, Tegan Maharaj, Evan D. Sherwin, S. Karthik Mukkavilli, Konrad P. Kording, Carla Gomes, Andrew Y. Ng, Demis Hassabis, John C. Platt15, Felix Creutzig, Jennifer Chayes, Yoshua Bengio6, "Tackling Climate Change with Machine Learning,"arXiv:1906.05433v1 [cs.CY] 10 June 2019.

[2] E. B. Abrahamsen, O. M. Brastein, B. Lie," Machine Learning in Python for Weather Forecast based on Freely Available Weather Data," Proceedings of The 59th Conference on Simulation and Modelling (SIMS 59), 26-28 September 2018, Oslo Metropolitan University, Norway.

[3] Y.Radhika and M.Shashi," Atmospheric Temperature Prediction using Support Vector Machines," International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009 1793-8201.

[4] Antonio S. Cofino˜ and Rafael Cano and Carmen Sordo and Jose M. Guti ´ errez ´," Bayesian Networks for Probabilistic Weather Prediction." in ECAI 2002. Proceedings of the 15th European Conference on Artificial Intelligence, IOS Press, 695 - 700 (2002).

[5] Mohsen Hayati, and Zahra Mohebi," Application of Artificial Neural Networks for Temperature Forecasting," World Academy of Science, Engineering and Technology International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering Vol:1, No:4, 2007.

[6] S Karmakar, M K Kowar, P Guhathakurta," Long-Range Monsoon Rainfall Pattern Recognition & Prediction for the Subdivision 'EPMB' Chhattisgarh Using Deterministic & Probabilistic Neural Network," 2009 Seventh International Conference on Advances in Pattern Recognition.

[7] Navin Sharma, Pranshu Sharma, David Irwin, and Prashant Shenoy," Predicting Solar Generation from Weather Forecasts Using Machine Learning," ©2011 IEEE.

[8] Deepti Gupta, Udayan Ghose," A Comparative Study of Classification Algorithms for Forecasting Rainfall," 2015 IEEE.

[9] T.R.V.Anandharajan, G.Abhishek Hariharan, K.K.Vignajeth, R.Jijendiran,Kushmita," WEATHER MONITORING USING ARTIFICIAL INTELLIGENCE," 2016 International Conference on Computational Intelligence and Networks.

[10] Shubham Madan, Praveen Kumar, Seema Rawat, Tanupriya Choudhury," Analysis of Weather Prediction using Machine Learning & Big Data," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE-2018) Paris, France 22-23 June 2018.

[11] Patricia de Rosnay, Nemesio Rodr´ıguez-Fernandez ´, Joaqu´ın Munoz-Sabater ˜, Clement Albergel ´, David Fairbairn, Heather Lawrence, Stephen English1 , Matthias Drusch and Yann Kerr," SMOS DATA ASSIMILATION FOR NUMERICAL WEATHER PREDICTION," ©2018 IEEE.

[12] Sue Ellen Haupt, Jim Cowie, Seth Linden, Tyler McCandless, Branko Kosovic, Stefano Alessandrini," Machine Learning for Applied Weather Prediction," 2018 IEEE 14th International Conference on e-Science.

[13] W. Enke, A. Spekat," Downscaling climate model outputs into local and regional weather elements by classification and regression," Vol. 8: 195-207,1997 Vol. 8: 195-207,1997.

[14] Dr. S. Santhosh Baboo, I.Kadar Shereef," An Efficient Weather Forecasting System using a Hybrid Neural Network SOFM–MLP," Volume 1, No.4, Nov-Dec 2010.

[15] Bjarne K. Hansen and Denis Riordan," Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory," 2001.

[16] Zahra Karevan, Siamak Mehrkanoon, Johan A.K. Suykens," Black-box modeling for temperature prediction in weather forecasting".

[17] Tomonobu Senjyu, Member, IEEE, Hitoshi Takara, Katsumi Uezato, and Toshihisa Funabashi, Senior Member, IEEE," One-Hour-Ahead Load Forecasting Using Neural Network," IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 17, NO. 1, FEBRUARY 2002.

[18] Ye Ding, Member, IEEE, Yanhua Li, Senior Member, IEEE, Ke Deng, Member, IEEE, Haoyu Tan, Member, IEEE, Mingxuan Yuan, Member, IEEE, Lionel M. Ni, Fellow, IEEE," Detecting and Analyzing Urban Regions with High Impact of Weather Change on Transport".10.1109/TBDATA.2016.2623320,IEEE.